```
Sequence length 2202
         10        20        30        40        50        60        70
GGAGGGCCTGAAGAGACAGGGAGGTTGTGCCAGGCTGGAGGAGGCTTGTCTTTCCGAAGCTGGAGAGGATCTTACGGGG
80       90        100       110       120       130       140       150
GTTCGCTTTTCCCTGCCTGGGAAGAATTTCCCCTGTGGTAGCAGCAGCAGCAGCAGCAGAAGCAGAAACAGCAGCAGCA
160       170       180       190       200       210       220       M   P   V
GCAACAGCAGCAGCAGCAGCAGCACCACCACCACCACTACCTCCTCTTCTGGGGCACAAGACAGA ATG CCT GTG    232
 L   E   R   Y   F   H   P   A   E   L   G   R   R   W   T   G   P   E   G   V
CTA GAG CGA TAT TTC CAC CCA GCA GAG CTA GGC AGG AGG TGG ACA GGC CCA GAA GGT GTG  292
 L   P   S   S   P   G   S   R   P   G   C   Q   Q   G   P   L   P   W   D   L
CTG CCC TCC TCC CCG GGA AGC CGG CCG GGG TGC CAG CAG GGG CCG CTG CCC TGG GAC TTG  352
 P   E   M   I   R   M   V   K   L   V   W   K   S   K   S   E   L   Q   A   T
CCA GAG ATG ATC AGG ATG GTA AAG CTG GTT TGG AAA TCC AAA AGT GAG CTG CAG GCG ACC  412
 K   Q   R   G   I   L   D   N   E   D   A   L   R   S   F   P   G   D   I   R
AAA CAG AGA GGC ATT CTG GAC AAT GAA GAT GCT CTC CGC AGC TTT CCA GGA GAT ATA CGA  472
 L   R   G   Q   T   G   V   R   A   E   R   R   G   S   Y   P   F   I   D   F
CTA AGG GGT CAG ACG GGG GTT CGT GCT GAA CGC CGT GGC TCC TAC CCA TTC ATT GAC TTC  532
 R   L   L   N   S   T   T   Y   S   G   E   I   G   T   K   K   K   V   K   R
CGC CTA CTT AAC AGT ACA ACA TAC TCA GGG GAG ATT GGC ACC AAG AAA AAG GTG AAA AGA  592
 L   L   S   F   Q   R   Y   F   H   A   S   R   L   L   R   G   I   I   P   Q
CTA TTA AGC TTT CAA AGA TAC TTC CAT GCA TCA AGG CTG CTT CGT GGA ATT ATA CCA CAA  652
 A   P   L   H   L   L   D   E   D   Y   L   G   Q   A   R   H   M   L   S   K
GCC CCT CTG CAC CTG CTG GAT GAA GAC TAC CTT GGA CAA GCA AGG CAT ATG CTC TCC AAA  712
 V   G   M   W   D   F   D   I   F   L   F   D   R   L   T   N   G   N   S   L
GTG GGA ATG TGG GAT TTT GAC ATT TTC TTG TTT GAT CGC TTG ACA AAT GGA AAC AGC CTG  772
 V   T   L   L   C   H   L   F   N   T   H   G   L   I   H   H   F   K   L   D
GTA ACA CTG TTG TGC CAC CTC TTC AAT ACC CAT GGA CTC ATT CAC CAT TTC AAG TTA GAT  832
 M   V   T   L   H   R   F   L   V   M   V   Q   E   D   Y   H   S   Q   N   P
ATG GTG ACC TTA CAC CGA TTT TTA GTC ATG GTT CAA GAA GAT TAC CAC AGC CAA AAC CCG  892
 Y   H   N   A   V   H   A   A   D   V   T   Q   A   M   H   C   Y   L   K   E
TAT CAC AAT GCT GTT CAC GCA GCC GAC GTC ACC CAG GCC ATG CAC TGC TAC CTG AAA GAG  952
 P   K   L   A   S   F   L   T   P   L   D   I   M   L   G   L   L   A   A   A
CCA AAG CTT GCC AGC TTC CTC ACG CCT CTG GAC ATC ATG CTT GGA CTG CTG GCT GCA GCA  1012
 A   H   D   V   D   H   P   G   V   N   Q   P   F   L   I   K   T   N   H   H
GCA CAC GAT GTG GAC CAC CCA GGG GTG AAC CAG CCA TTT TTG ATA AAA ACT AAC CAC CAT  1072
 L   G   A   L   Y   Q   N   M   S   V   L   E   N   H   H   W   R   S   T   I
CTT GCA AAC CTA TAT CAG AAT ATG TCT GTG CTG GAG AAT CAT CAC TGG CGA TCT ACA ATT  1132
 G   M   L   R   E   S   R   L   L   A   H   L   P   K   E   M   T   Q   D   I
GGC ATG CTT CGA GAA TCA AGG CTT CTT GCT CAT TTG CCA AAG GAA ATG ACA CAG GAT ATT  1192
 E   Q   Q   L   G   S   L   I   L   A   T   D   I   N   R   Q   N   E   F   L
GAA CAG CAG CTG GGC TCC TTG ATC TTG GCA ACA GAC ATC AAC AGG CAG AAT GAA TTT TTG  1252
```

FIG. 1A.

T R L K A H L H N K D L R L E D A Q D R
ACC AGA TTG AAA GCT CAC CTC CAC AAT AAA GAC TTA AGA CTG GAG GAT GCA CAG GAC AGG 1312
H F M L Q I A W K C A D I C N P C R I W
CAC TTT ATG CTT CAG ATC GCT TGG AAG TGT GCT GAC ATT TGC AAT CCT TGT AGA ATC TGG 1372
E M S K Q W S E R V C E E F Y R Q G E L
GAG ATG AGC AAG CAG TGG AGT GAA AGG GTC TGT GAA GAA TTC TAC AGG CAA GGT GAA CTT 1432
E Q K F E L E I S P L C N Q Q K D S I P
GAA CAG AAA TTT GAA CTG GAA ATC AGT CCT CTT TGT AAT CAA CAG AAA GAT TCC ATC CCT 1492
S I Q I G F M S Y I V E P L F R E W A H
AGT ATA CAA ATT GGT TTC ATG AGC TAC ATC GTG GAG CCG CTC TTC CGG GAA TGG GCC CAT 1552
F T G N S T L S E N M L G H L A H N K A
TTC ACG GGT AAC AGC ACC CTG TCG GAG AAC ATG CTG GGC CAC CTC GCA CAC AAC AAG GCC 1612
Q W K S L L P R Q H R S R G S S G S G P
CAG TGG AAG AGC CTG TTG CCC AGG CAG CAC AGA AGC AGG GGC AGC AGT GGC AGC GGG CCT 1672
D H D H A G Q G T E S E E Q E G D S P *
GAC CAC GAC CAC GCA GGC CAA GGG ACT GAG AGC GAG GAG CAG GAA GGC GAC AGC CCC TAG 1732

GGGCCGGCCCAACTTAGACGCGGCTCTCCTCCGGCAGGGCCCCCAGAGGGCAGAAGCAGCGTGGAGGGGCCCTCACGCA

GCAGCCCAGCCACTTTCTGAGTGTTGTCCTGGGGCTCTTTGGAACGCCATCTTCCTCCCACTTACCTGCCTCCCCTCCT

TTTCGCAAATGTACAGAAGCCATTTGTCACCTCAGCATTCGCTGCCGAAATGAGCAACTCCATTCAGTAAGCTGGGAGC

TGATCCCACGGGCAGGCTCTCCCTGCTCCAGGAGAAGACTAGGAGGAAGAATGAGGTGCTCCTGCCGTGTCCGCCTTGT

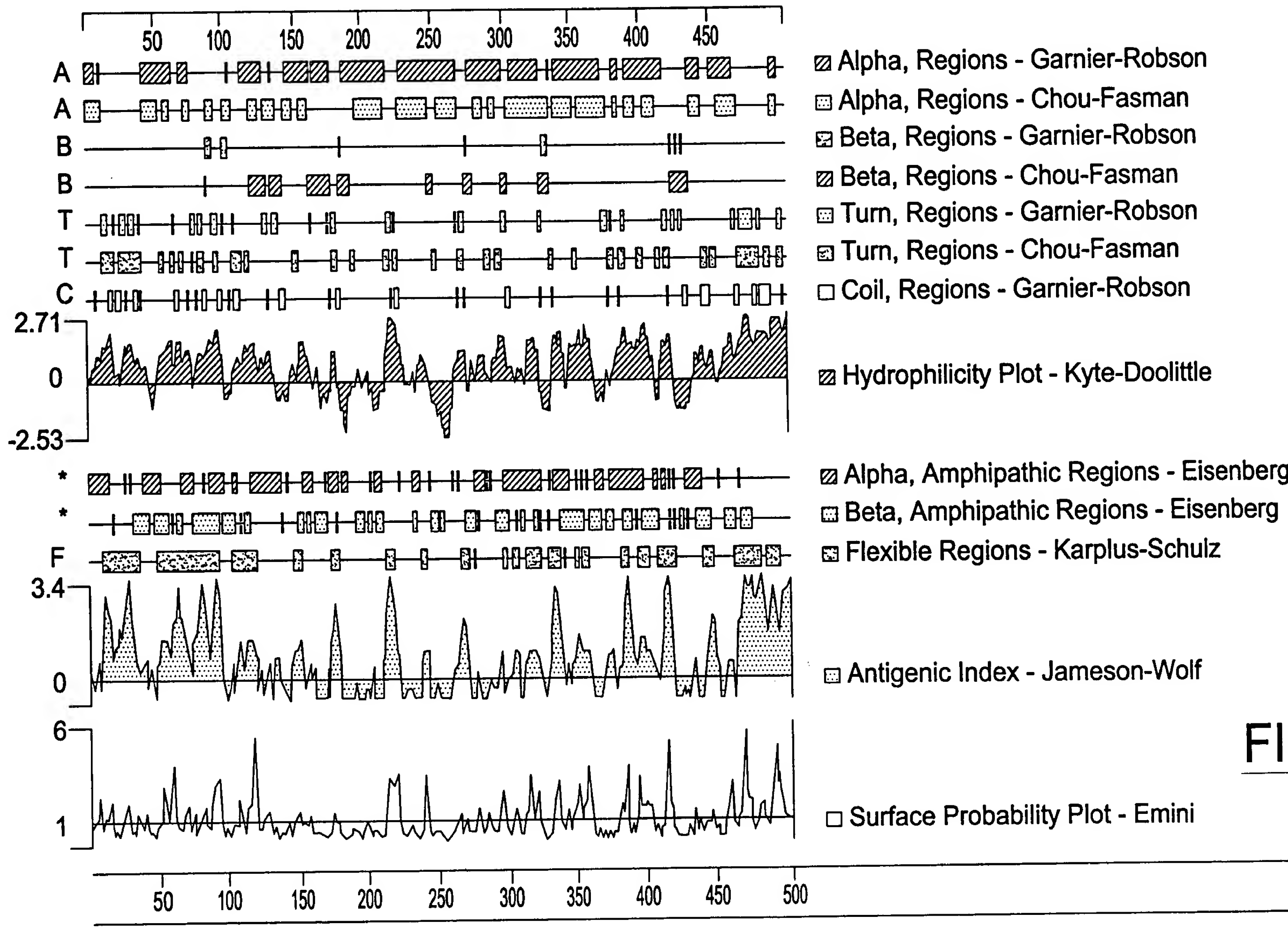
TCCGGGTCGCACTGGAACAGGCAGCAATTCCTAAGTCCGGAGCGTTTGAGCGTTTGCTATCTGACTGCTGATCTGCGTG

ACAGAAACACCAGCATATTTGCAACGCCAAGGATATTGGTCTTAAAGTGCAAGAACACAAATGAGAGTGTGAAGA

FIG. 1B.

```
PDEase   PF00233 3'5'-cyclic nucleotide phosphodieste   211.4   3.3e-73   1

Parsed for domains:
Model    Domain  seq-f seq-t    hmm-f hmm-t      score  E-value
-------- ------- ----- -----    ----- -----      -----  -------
PDEase     1/1    224   462 ..      1   279 []   211.4  3.3e-73

Alignments of top-scoring domains:
PDEase:  domain 1 of 1, from 224 to 462:  scor 211.4, E = 3.3e-73
                     *->YHNwiHAfdVtQtthlLlltlaleryLtdlEvLalvfAAaiHDvDHr
                        YHN  HA+dVtQ++h+ l+ ++l   Lt+l + + + AAa HDvDH+
   sequence27   224     YHNAVHAADVTQAMHCYLKEPKLASFLTPLDIMLGLLAAAAHDVDHP 270

                     GTnNsFqinsLqkSeLAlLYndegSVLEnHHlaqafkLLqdEecnIfqNL
                     G+n  F+i++  +  LA LY +  SVLEnHH  +++ +L+  e    + +L
   sequence27   271  GVNQPFLIKT--NHHLANLYQNM-SVLENHHWRSTIGMLR--ESRLLAHL 315

                     skkdfrtlrdlviealLAATDmslHlqklkdlktmveqkkvyetgvEWtqY
                                                                       +
                     +k   +++        ILATD+        l++lk  +  k
   sequence27   316  PKEMTQDIEQQLGSLILATDINRQNEFFLTRLKAHLHNK----------DL 355

                     lldnythkillllsllmtAADLSnpTKpwslskRwAelimeEFFeQGDlEr
                      l + +++    l+    ++AD+ np  +w+ sk+w e++ eEF++QG lE
   sequence27   356  RLEDAQDRHFMLQIAWKCADICNPCRIWEMSKQWSERVCEEFYRQGELEQ 405

                     elGldrpspmcDRtsAayvpksQvgFidfIvePvfklLadvvekGRttse
                       l   +sp+c   + +++p  Q+gF+ +IveP+f ++a ++
   sequence27   406  KFELE-ISPLCNQQK-DSIPSIQEGFMSYIVEPLFREWAHFTGN------ 447

                     aiDanhLCWvaLDeevRnddiapllDriednR<-*
                                           +  +l +++ n+
   sequence27   448  -------------------STLSENMLGHLAHNK
```

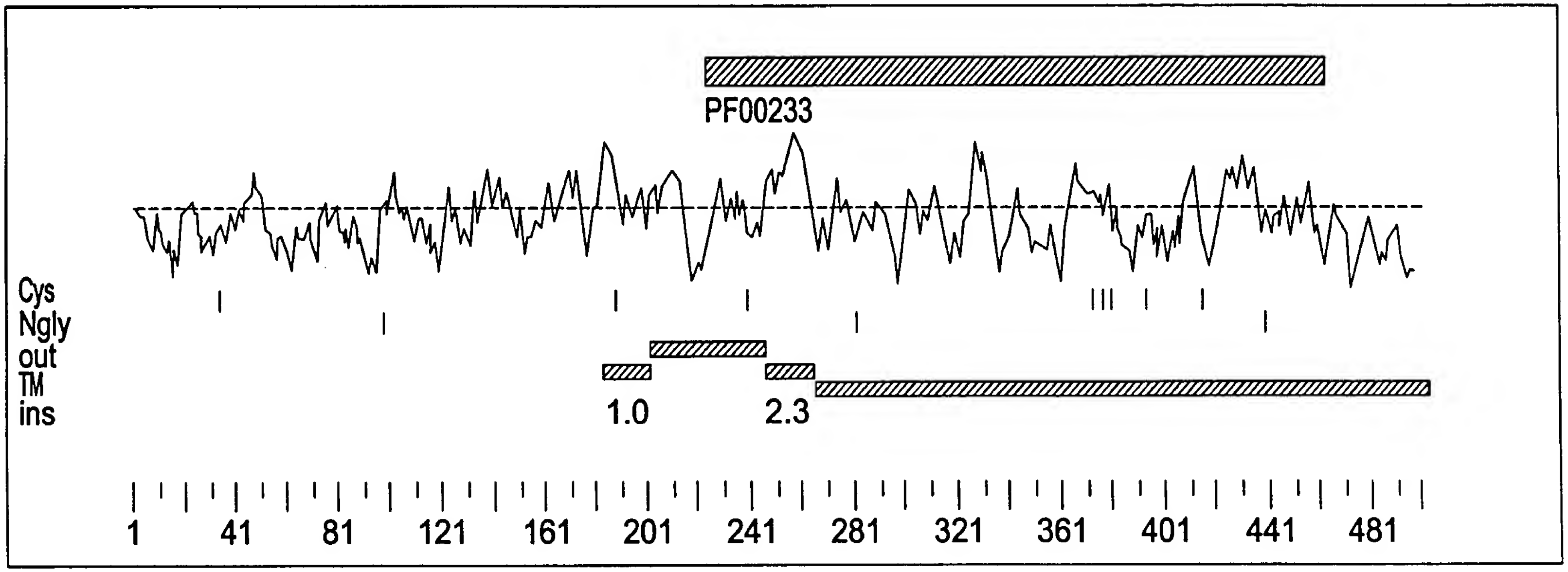
FIG. 2.



50 100 150 200 250 300 350 400 450
A
A
B
B
T
T
C
2.71
0
-2.53
*
*
F
3.4
0
6
1
50 100 150 200 250 300 350 400 450 500
Alpha, Regions - Garnier-Robson
Alpha, Regions - Chou-Fasman
Beta, Regions - Garnier-Robson
Beta, Regions - Chou-Fasman
Turn, Regions - Garnier-Robson
Turn, Regions - Chou-Fasman
Coil, Regions - Garnier-Robson
Hydrophilicity Plot - Kyte-Doolittle
Alpha, Amphipathic Regions - Eisenberg
Beta, Amphipathic Regions - Eisenberg
Flexible Regions - Karplus-Schulz
Antigenic Index - Jameson-Wolf
Surface Probability Plot - Emini

FIG. 3.

Analysis of sequence2741 (502 aa)




```
>sequence2741
MPVLERYFHPAELGRRWTGPEGVLPSSPGSRPGCQQGPLPWDLPEMIRMVKLVWKSKSEL
QATKQRGILDNEDALRSFPGDIRLRGQTGVRAERRGSYPFIDFRLLNSTTYSGEIGTKKK
VKRLLSFQRYFHASRLLRGIIPQAPLHLLDEDYLGQARHMLSKVGMWDFDIFLFDRLTNG
NSLVTLLCHLFNTHGLIHHFKLDMVTLHRFLVMVQEDYHSQNPYHNAVHAADVTQAMHCY
LKEPKLASFLTPLDIMLGLLAAAAHDVDHPGVNQPFLIKTNHHLANLYQNMSVLENHHWR
STIGMLRESRLLAHLPKEMTQDIEQQLGSLILATDINRQNEFLTRLKAHLHHKDLRLEDA
QDRHFMLQIAWKCADICNPCRIWEMSKQWSERVCEEFYRQGELEQKFELEISPLCNQQKD
SIPSIQIGFMSYIVEPLFREWAHFTGNSTLSENMLGHLAHNKAQWKSLLPRQHRSRGSSG
SGPDHDHAGQGTESEEQEGDSP
```

FIG. 4.

Transmembrane Segments Predicted by MEMSAT

Start	End	Orient	Score
183	200	ins-->out	1.0
246	264	out-->ins	2.3

```
>sequence2741
MPVLERYFHPAELGRRWTGPEGVLPSSPGSRPGCQQGPLPWDLPEMIRMVKLVWKSKSEL
QATKQRGILDNEDALRSFPGDIRLRGQTGVRAERRGSYPFIDFRLLNSTTYSGEIGTKKK
VKRLLSFQRYFHASRLLRGIIPQAPLHLLDEDYLGQARHMLSKVGMWDFDIFLFDRLTNG
NSLVTLLCHLFNTHGLIHHFKLDMVTLHRFLVMVQEDYHSQNPYHNAVHAADVTQAMHCY
LKEPKLASFLTPLDIMGLLAAAAHDVDHPGVNQPFLIKTNHHLANLYQNMSVLENHHWR
STIGMLRESRLLAHLPKEMTQDIEQQLGSLILATDINRQNEFLTRLKAHHLHNKDLRLEDA
QDRHFMLQIAWKCADICNPCRIWEMSKQWSERVCEEFYRQGELEQKFELEISPLCNQQKD
SIPSIQIGFMSYIVEPLFREWAHFTGNSTLSENMLGHLAHNKAQWKSLLPRQHRSRGSSG
SGPDHDHAGQGTESEEQEGDSP
```

Prosite Pattern Matches for sequence2741

>PS00001/PDOC00001/ASB_GLYCOSYLATION N-glycosylation site.

Query:	107	NSTT	110
Query:	290	NMSV	293
Query:	447	NSTL	450

FIG. 5A.

>PS00002/PDOC00002/GLYCOSAMINOGLYCAN Glycosaminoglycan attachment site.
RU Additional rules:
RU There must be at least two acidic amino acids (Glu or Asp) from -2 to
RU -4 relative to the serine.

Query: 479 SGSG 482

>PS00004/PDOC00004/CAMP_PHOSPHO_SITE cAMP- and cGMP-dependent protein kinase phosphorylation site.

Query: 15 SGSG 18
Query: 94 RRGS 97

>PS00005/PDOC00005/PKC_PHOSPHO_SITE Protein kinase C phosphorylation site.

Query: 117 TKK 119
Query: 390 SER 392

>PS00006/PDOC00006/CK2_PHOSPHO_SITE Casein kinase II phosphorylation site.

Query: 18 TGPE 21
Query: 56 SKSE 59
Query: 251 TPLD 254
Query: 292 SVLE 295
Query: 449 TLSE 452
Query: 481 SGPD 484
Query: 492 TESE 495

>PS00007/PDOC00007/TYR_PHOSPHO_SITE Tyrosine kinase phosphorylation site.

Query: 392 RVCEEFY 398

>PS00008/PDOC00008/MYRISTYL N-myristoylation site.

Query: 22 GVLPSS 27
Query: 29 GSRPGC 34
Query: 67 GILDNE 72
Query: 258 GLLAAA 263
Query: 477 GSSGSG 482

>PS00009/PDOC00009/AMIDATION Amidation site.

Query: 13 LGRR 16

FIG. 5B.

22025c12, 3336 bases, 672 check
[Strand)

```
1     GAGGGCCTGA AGACACAGGG AGGTTGTGCC AGGCTGGAGG

41    AGGCTTGTCT TTCCGAAGCT GGAGAGGATC TTACGGGGGT

81    TCGCTTTTCC CTGCCTGGGA AGAATTTCCC CTGTGGTAGC

121   AGCAGCAGCA GCAGCAGAAG CAGAAACAGC AGCAGCAGCA

161   ACAGCAGCAG CAGCAGCAGC ACCACCACCA CCACTACCTC

201   CTCTTCTGGG GCACAAGACA GAATGCCTGT GCTAGAGCGC
                                 METProVal LeuGluArg
241   TATTTCCACC CAGCAGAGCT AGGCAGGAGG TGGACAGGCC
      ThyPheHisPro AlaGluLeu GlyArgArg TryThrGlyP
281   CAGAAGGTGT GCTGCCCTCC TCCCCGGGAA GCCGGCCGGG
      roGluGlyVal LeuProSer SerProGlySer ArgProGl
321   GTGCCAGCAG GGGCCGCTGC CCTGGGACTT GCCAGAGATG
      yCysGlnGln GlyProLeuPro TrpAspLeu ProGluMET
361   ATCAGGATGG TAAAGCTGGT TTGGAAATCC AAAAGTGAGC
      IleArgMETVal LysLeuVal TrpLysSer LysSerGluL
401   TGCAGGCGAC CAAACAGAGA GGCATTCTGG ACAATGAAGA
      euGlnAlaThr LysGlnArg GlyIleLeuAsp AsnGluAs
441   TGCTCTCCGC AGCTTTCCAG GAGATATACG ACTAAGGGGT
      pAlaLeuArg SerPheProGly AspIleArg LeuArgGly
481   CAGACGGGGG TTCGTGCTGA ACGCCGTGGC TCCTACCCAT
      GlnThrGlyVal ArgAlaGlu ArgArgGly SerTyrProP
521   TCATTGACTT CCGCCTACTT AACAGTACAA CATACTCAGG
      heIleAspPhe ArgLeuLeu AsnSerThrThr TyrSerGl
561   GGAGATTGGC ACCAAGAAAA AGGTGAAAAG ACTATTAAGC
      yGluIleGly ThrLysLysLys ValLysArg LeuLeuSer
```

FIG. 6A.

22025c12, 3336 bases, 672 check
[Strand]

601 TTTCAAAGAT ACTTCCATGC ATCAAGGCTG CTTCGTGGAA
PheGlnArgTyr PheHisAla SerArgLeu LeuArgGlyI
641 TTATACCACA AGCCCCTCTG CACCTGCTGG ATGAAGACTA
leIleProGln AlaProLeu HisLeuLeuAsp GluAspTy
681 CCTTGGACAA GCAAGGCATA TGCTCTCCAA AGTGGGAATG
rLeuGlyGln AlaArgHisMET LeuSerLys ValGlyMET
721 TGGGATTTTG ACATTTTCTT GTTTGATCGC TTGACAAATG
TrpAspPheAsp IlePheLeu PheAspArg LeuThrAsnG
761 GAAACAGCCT GGTAACACTG TTGTGCCACC TCTTCAATAC
lyAsnSerLeu ValThrLeu LeuCysHisLeu PheAsnTh
801 CCATGGACTC ATTCACCATT TCAAGTTAGA TATGGTGACC
rHisGlyLeu IleHisHisPhe LysLeuAsp METValThr
841 TTACACCGAT TTTTAGTCAT GGTTCAAGAA GATTACCACA
LeuHisArgPhe LeuValMET ValGlnGlu AspTyrHisS
881 GCCAAAACCC GTATCACAAT GCTGTTCACG CAGCCGACGT
erGlnAsnPro TyrHisAsn AlaValHisAla AlaAspVa
921 CACCCAGGCC ATGCACTGCT ACCTGAAAGA GCCAAAGCTT
lThrGlnAla METHisCysTyr LeuLysGlu ProLysLeu
961 GCCAGCTTCC TCACGCCTCT GGACATCATG CTTGGACTGC
AlaSerPheLeu ThrProLeu AspIleMET LeuGlyLeuL
1001 TGGCTGCAGC AGCACACGAT GTGGACCACC CAGGGGTGAA
euAlaAlaAla AlaHisAsp ValAspHisPro GlyValAs
1041 CCAGCCATTT TTGATAAAAA CTAACCACCA TCTTGCAAAC
nGlnProPhe LeuIleLysThr AsnHisHis LeuAlaAsn
1081 CTATATCAGA ATATGTCTGT GCTGGAGAAT CATCACTGGC
LeuTyrGlnAsn METSerVal LeuGluAsn HisHisTrpA
1121 GATCTACAAT TGGCATGCTT CGAGAATCAA GGCTTCTTGC
rgSerThrIle GlyMETLeu ArgGluSerArg LeuLeuAl
1161 TCATTTGCCA AAGGAAATGA CGTAAGTGCT GCCGAGATGA
aHisLeuPro LysGluMETThr STP

FIG. 6B.

22025c12, 3336 bases, 672 check
[Strand]

1201 AACATACTGA TGTGCATGCA GTAAAGATAA GCCACTTTCT

1241 CTAGGGCAGG CTTGGGACCT TTTGCGTGAA TGGCAGAGAG

1281 CCCCCCGCT GTACTTCCTG CCTGCACTGA GCTGTCTATC

1321 AGAGGAGATT TGGTGTCAGT TACAGCAACC CAGAAACCAA

1361 AATCTCTCTG TGTGCTTTGA AAGGGCCTTG CAGAGTCAAT

1401 GACCTACAGT CAGGAAAAGG GATAATAAAC AGCTCTCAGT

1441 TTTCACACGC TTCAGTATCA GTGCTCGACT TTGCCAAATT

1481 CCCGACCTTT AGTTTAGCAA AATTGTCCTT CCATGTAGCT

1521 CCAAATAGTA AATATTTATC AAGAAGGAAC CCAGGCATTC

1561 TAAAGCTAGA GTTCAAAAAA GTATATTTTG TAATTGCTAG

1601 TCTCAGCAAA AATAGAAGTC AGAAATTCTT TTCTAAAATG

1641 TCTTTTGCTA AGTAATTGAA ATGGCCCTAG CATTTTTTTC

1681 ACCAATTAAT TTACCTTACG TCTCTTGCAC TTTAAACAGA

1721 AGGGGAGACA CTCATTTTCT GGTTCACTAT TTGATAGCCA

1761 TGGTATGTAG GCTGAGTCCC ACTAAATCTG AGGCCATTGT

FIG. 6C.

22025c12, 3336 bases, 672 check
[Strand]

1801 TTCATTTTCC TGGTGGCCCC AAGTTAGCTG CTAATACTGT

1841 CTTCCAAGGC CACCATTAAT TCTGATCTGT TTAATGAACA

1881 CGTGCAGAAC CCAAGAAACC TAGGTGAAAA GAGTACATAG

1921 ATTGCTGTAC CCTTCTTCAA GACAAGCACA TAACTTGAGG

1961 TCAAGGACCA AGTGCTGTCT CCCAACTGAA CAAGCAGTAT

2001 ACTCTGGGTT GTGGATTGAT TCCTGGCCCT CTGATTTGAT

2041 CTCATGCTGT TTCCTAGCAC CCAGAGGAAT GTGAAATTTG

2081 CAGGAGGAAT TTCAGTTCTG ATAAATTTTT ACTCCCTGGA

2121 ACTAAATAAA ACCAGTTCTC GTGCATGGAA TAAAAACTTA

2161 TGCCTCTTAC TAGAATAATA AATTGCAAAG ATTGAAAGAA

2201 TTAAATGCAA AAAGAACTAA AAACTAGAGC AAAAGATCAA

2241 GTGAGAAGAA GAAAAGAGGA GGTAAGGAGA GAGACAAGGA

2281 AGAAAGAAGG AGAAGGAAAG GAAGAATAGT GAGGACAGGA

2321 AAGAAGAAAA TGCAAGGGAA ATGGGAAAGG ACTCTGGGGT

2361 GACCAGACTT CTCCTGGTCA GTACCTGCAT TCATCCTGTT

FIG. 6D.

22025c12, 3336 bases, 672 check
[Strand]

2401 TGTTACTCAA TATTTCTTTC CTAAAATATT CATTTCACAT

2441 CTATGGATTC CAATGAAAAA TATATTTTTA TGTGTCTTTG

2481 TGGAACACAG TGTTATAAAT TGTTTTTGCC AGAAGAATAA

2521 TTGTTATACA ATAATATATG TGAAAACTTT ATTACAAAAG

2561 CCATTATCAT AATCATTATT ATTCCTTCTA TCACAGGTAA

2601 ATGCTTTAAT GTCATTTTTC TGATTTTAAA AGTAGGGCAG

2641 GTTAATTGTA GAAAGTAAGG AAAATTCAGG AAAGTGTTAG

2681 TTTGAACTAT GTGAAGTTGC TCTTTTTAAG GGCCAAAAAC

2721 AGGAGACTTT TAGCACTTTC ATATGTTTCA GCTTGATATG

2761 AAAGAGAAAA CTGAAACTGC TAGTAATCCT GCCATCCAGG

2801 TATAGTTCAT GTTAACCTGG CTAGTTTATT TTCTTTTAGT

2841 CTTTTTTCAA TACAAACTTA TTTTAACAAA ATATGATTAN

2881 ATTTGGGGAA CTTATTTTAC AGTTTACGTC CTGAAATTTT

2921 TTATTTACAA TAAAGACTTT TTTCCAAATC ATTAAACCTG

2961 TTAAATTAAA ATGATTTTGT CAGCCGTATG GCATTATTGT

FIG. 6E.

22025c12, 3336 bases, 672 check
[Strand]

3001 ATACCACTAC TGCCTTTCAT TTGGAATTCA AATGGTTTCC

3041 AATATCCCAA ACTTTGATAC TCTGTTTTCT CAGGAAGTAT

3081 TTGTAGATAA AAATTATTGG TCAGAAAGGT CTGAACTTTT

3121 AAGTTTCTTG TATATTATCC AGTTGTTCTT CTAAAAGGCT

3161 GTATCTACCT GTATTCCAAC TGATGGATTG TAAGAAAATG

3201 TACCAATGTA CCATCACCAA AATTGAGTTT ATTTTTATCT

3241 TTTTAAAATA TTTGCAAATT TGACATATAT GTATGTATAT

3281 ACACAAATAT ATATGTAAAG TGGTTTTCAT TAAATTAGTA

3321 TGCATCCTTT ACTTAC

FIG. 6F.

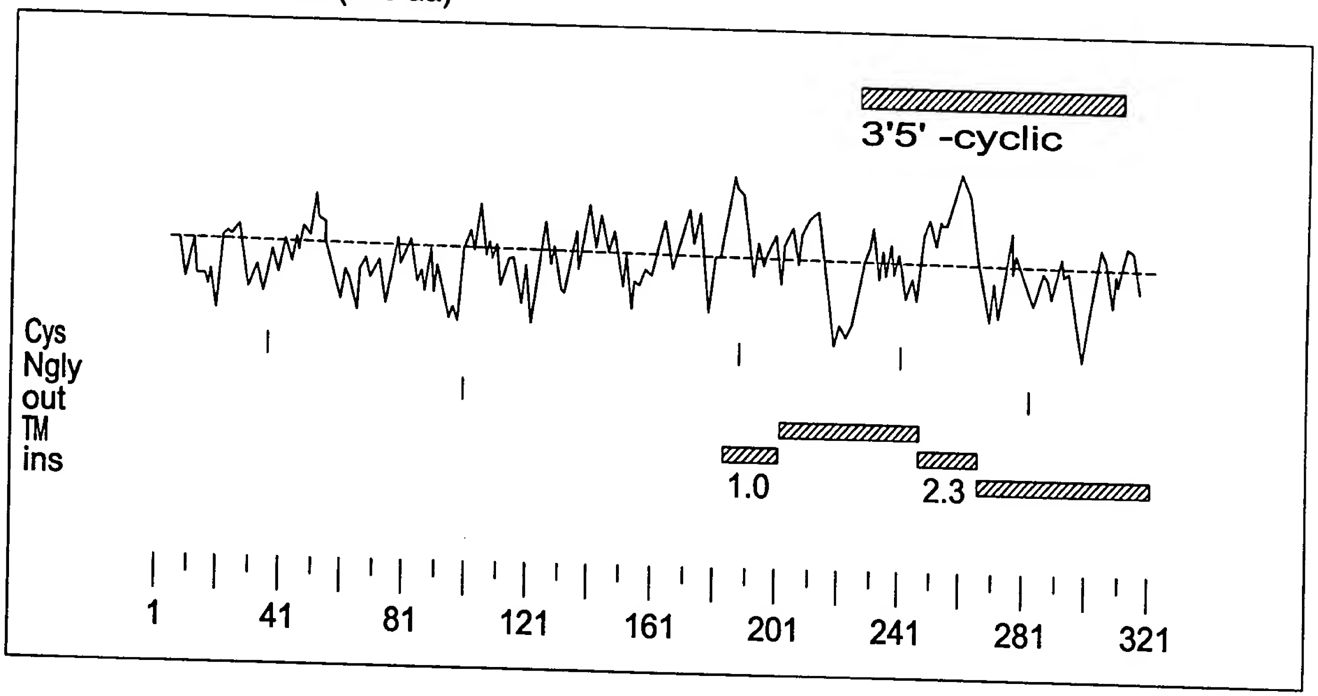
Protein Family / Domain Matches, HMMer version 2

```
-----------------------------------------------------------------
 Query:   22025short
Scores for sequence family classification (score includes all domains):
Model    Description                                      Score    E-value  N
-------- -----------                                      -----    -------  ---
PDEase   3'5' -cyclic nucleotide phosphodiesterase        110.8    3.8e-38   1
Parsed for domains:
Model    Domain  seq-f seq-t    hmm-f hmm-t      score  E-value
-------  ------- ----- -----    ----- -----      -----  -------
PDEase     1/1    224   308 ..     1    88 [.    110.8  3.8e-38
Alignments of top-scoring domains:
PDEase:  domain 1 of 1, from 224 to 308:  score 110.8, E = 3.8e-38
                   *->YHNwiHAfdVtQtthlLlltalaleryLtdlEvLalvfAAaiHDvDHr
                      YHN  HAA+dVtQ++h+ l+ ++l   Lt+; + + + AAa HDvDH+
  22025short   224    YHNAVHAADVTQAMHCYLKEPKLASFLTPLDIMLGLLAAAAHDVDHP 270

                   GTnNsFqinsLqkSeLAlLYndegSVLEnHHlaqafkLLqd<-*
                   G+n  F+i++  +  LA LY +  SVLEnHH  +++ +L++
  22025short   271 GVNQPFLIKT--NHHLANLYQNM-SVLENHHWRSTIGMLRE    308
```

FIG. 7.

Back to orfanal.cgi

Analysis of 22025short (320 aa)




```
>22025short
MPVLERYFHPAELGRRWTGPEGVLPSSPGSRPGCQQGPLPWDLPEMIRMVKLVWKSKSEL
QATKQRGILDNEDALRSFPGDIRLRGQTGVRAERRGSYPFIDFRLLNSTTYSGEIGTKKK
VKRLLSFQRYFHASRLLRGIIPQAPLHLLDEDYLGQARHMLSKVGMWDFDIFLFDRLTNG
NSLVTLLCHLFNTHGLIHHFKLDMVTLHRFLVMVQEDYHSQNPYHNAVHAADVTQAMHCY
LKEPKLASFLTPLDIMLGLLAAAAHDVDHPGVNQPFLIKTNHHLANLYQNMSVLENHHWR
STIGMLRESRLLAHLPKEMT
```

FIG. 8.

Prosite Pattern Matches for 22025short

Prosite version: Release 12.2 of February 1995

>PS00001/PDOC00001/ASN_GLYCOSYLATION N-glycosylation site.

Query:	107	NSTT	110
Query:	290	NMSV	293

>PS00004/PDOC00004/CAMP_PHOSPHO_SITE cAMP- and cGMP-dependent protein kinase phosphorylation site.

Query:	15	RRWT	18
Query:	94	RRGS	97

>PS00005/PDOC00005/PKC_PHOSPHO_SITE Protein kinase C phosphorylation site.

Query:	117	TKK	119

>PS00006/PDOC00006/CK2_PHOSPHO_SITE Casein kinase II phosphorylation site.

Query:	18	RRGS	21
Query:	56	RRGS	59
Query:	251	TKK	254
Query:	292	TKK	295

>PS00008/PDOC00008/MYRISTYL N-myristoylation site.

Query:	22	GVLPSS	27
Query:	29	GSRPGC	34
Query:	67	GILDNE	72
Query:	258	GLLAAA	263

>PS00009/PDOC00009/AMIDATION Amidation site.

Query:	13	LGRR	16

FIG. 9.